



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 08:34 AM UTC

PDB ID : 5GT7 / pdb_00005gt7
Title : Crystal Structure of Arg-bound CASTOR1
Authors : Guo, L.; Deng, D.
Deposited on : 2016-08-18
Resolution : 2.05 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

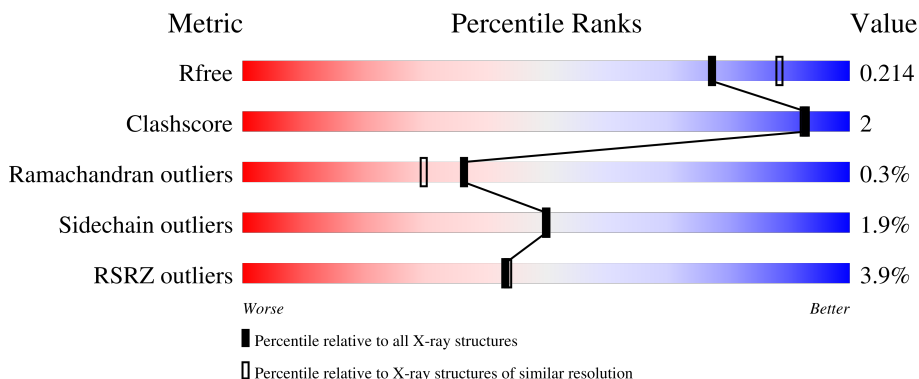
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.05 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	2260 (2.04-2.04)
Clashscore	190562	2333 (2.04-2.04)
Ramachandran outliers	187476	2318 (2.04-2.04)
Sidechain outliers	187428	2318 (2.04-2.04)
RSRZ outliers	180081	2260 (2.04-2.04)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	335	 2% 88% 6% 4%
1	B	335	 3% 84% 6% 7%
1	C	335	 5% 88% 6% 2%
1	D	335	 4% 82% 6% 10%

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 19732 atoms, of which 9594 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called GATS-like protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	305	4767	1542	2384	392	440	9	0	2	0
1	B	296	4666	1510	2335	381	432	8	0	4	0
1	C	318	4946	1596	2467	409	465	9	0	3	0
1	D	300	4695	1515	2345	389	438	8	0	2	0

There are 48 discrepancies between the modelled and reference sequences:

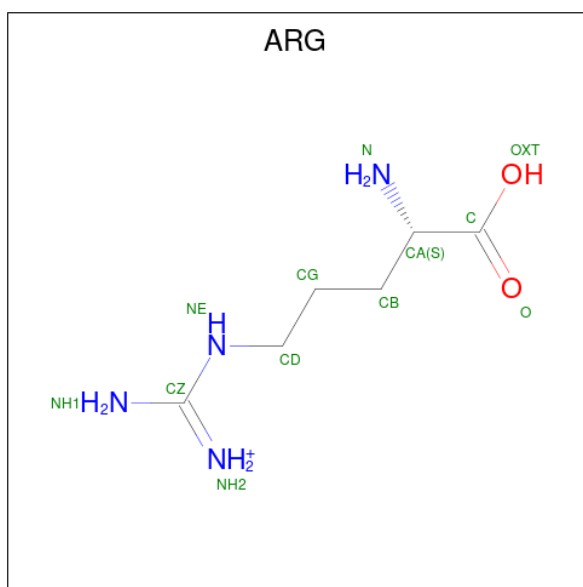
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	ALA	-	expression tag	UNP Q8WTX7
A	-4	GLY	-	expression tag	UNP Q8WTX7
A	-3	SER	-	expression tag	UNP Q8WTX7
A	-2	GLY	-	expression tag	UNP Q8WTX7
A	-1	HIS	-	expression tag	UNP Q8WTX7
A	0	MET	-	expression tag	UNP Q8WTX7
A	324	GLN	-	expression tag	UNP Q8WTX7
A	325	GLU	-	expression tag	UNP Q8WTX7
A	326	GLY	-	expression tag	UNP Q8WTX7
A	327	LEU	-	expression tag	UNP Q8WTX7
A	328	ALA	-	expression tag	UNP Q8WTX7
A	329	SER	-	expression tag	UNP Q8WTX7
B	-5	ALA	-	expression tag	UNP Q8WTX7
B	-4	GLY	-	expression tag	UNP Q8WTX7
B	-3	SER	-	expression tag	UNP Q8WTX7
B	-2	GLY	-	expression tag	UNP Q8WTX7
B	-1	HIS	-	expression tag	UNP Q8WTX7
B	0	MET	-	expression tag	UNP Q8WTX7
B	324	GLN	-	expression tag	UNP Q8WTX7
B	325	GLU	-	expression tag	UNP Q8WTX7
B	326	GLY	-	expression tag	UNP Q8WTX7

Continued on next page...

Continued from previous page...

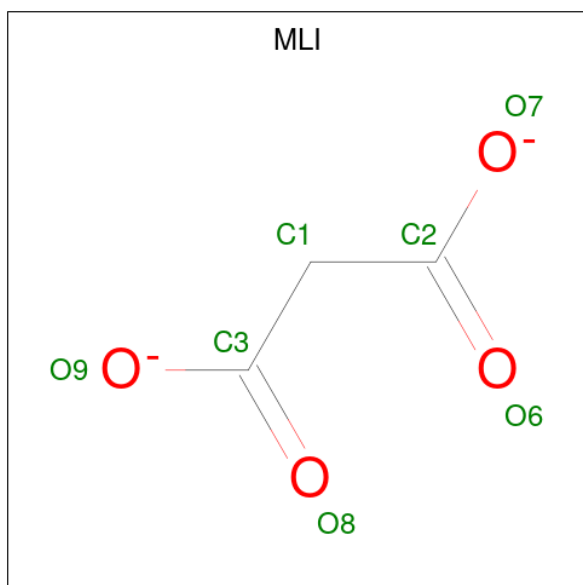
Chain	Residue	Modelled	Actual	Comment	Reference
B	327	LEU	-	expression tag	UNP Q8WTX7
B	328	ALA	-	expression tag	UNP Q8WTX7
B	329	SER	-	expression tag	UNP Q8WTX7
C	-5	ALA	-	expression tag	UNP Q8WTX7
C	-4	GLY	-	expression tag	UNP Q8WTX7
C	-3	SER	-	expression tag	UNP Q8WTX7
C	-2	GLY	-	expression tag	UNP Q8WTX7
C	-1	HIS	-	expression tag	UNP Q8WTX7
C	0	MET	-	expression tag	UNP Q8WTX7
C	324	GLN	-	expression tag	UNP Q8WTX7
C	325	GLU	-	expression tag	UNP Q8WTX7
C	326	GLY	-	expression tag	UNP Q8WTX7
C	327	LEU	-	expression tag	UNP Q8WTX7
C	328	ALA	-	expression tag	UNP Q8WTX7
C	329	SER	-	expression tag	UNP Q8WTX7
D	-5	ALA	-	expression tag	UNP Q8WTX7
D	-4	GLY	-	expression tag	UNP Q8WTX7
D	-3	SER	-	expression tag	UNP Q8WTX7
D	-2	GLY	-	expression tag	UNP Q8WTX7
D	-1	HIS	-	expression tag	UNP Q8WTX7
D	0	MET	-	expression tag	UNP Q8WTX7
D	324	GLN	-	expression tag	UNP Q8WTX7
D	325	GLU	-	expression tag	UNP Q8WTX7
D	326	GLY	-	expression tag	UNP Q8WTX7
D	327	LEU	-	expression tag	UNP Q8WTX7
D	328	ALA	-	expression tag	UNP Q8WTX7
D	329	SER	-	expression tag	UNP Q8WTX7

- Molecule 2 is ARGinine (CCD ID: ARG) (formula: C₆H₁₅N₄O₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	A	1	Total	C	H	N	O	0	0
			27	6	15	4	2		
2	B	1	Total	C	H	N	O	0	0
			24	6	12	4	2		
2	C	1	Total	C	H	N	O	0	0
			24	6	12	4	2		
2	D	1	Total	C	H	N	O	0	0
			24	6	12	4	2		

- Molecule 3 is MALONATE ION (CCD ID: MLI) (formula: $C_3H_2O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	A	1	Total	C	H	O	0	0
			9	3	2	4		
3	C	1	Total	C	H	O	0	0
			9	3	2	4		
3	D	1	Total	C	H	O	0	0
			9	3	2	4		
3	D	1	Total	C	H	O	0	0
			9	3	2	4		

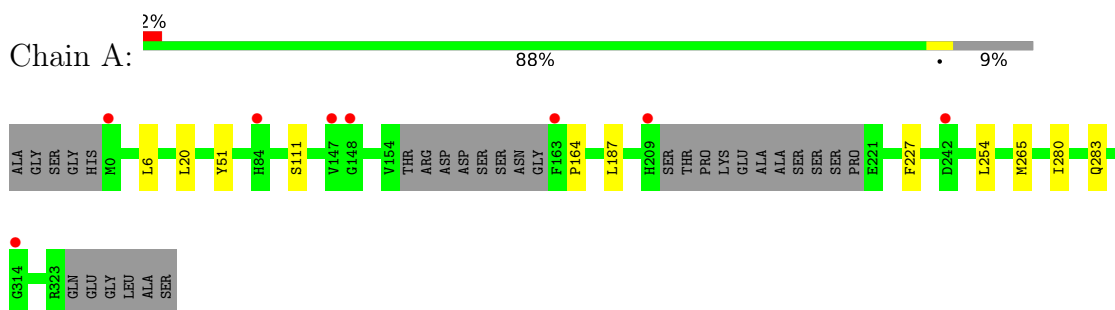
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	115	Total	O	0	0
			115	115		
4	B	104	Total	O	0	0
			104	104		
4	C	158	Total	O	0	0
			158	158		
4	D	128	Total	O	0	0
			128	128		

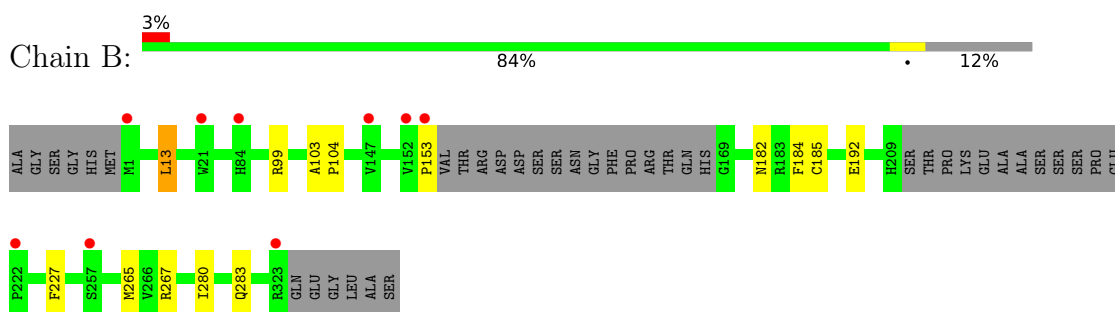
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

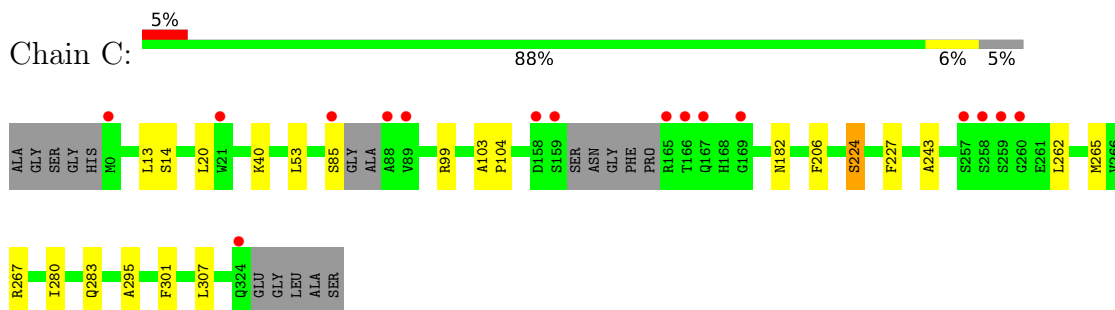
- Molecule 1: GATS-like protein 3



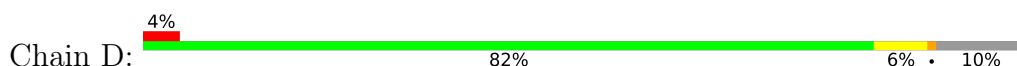
- Molecule 1: GATS-like protein 3

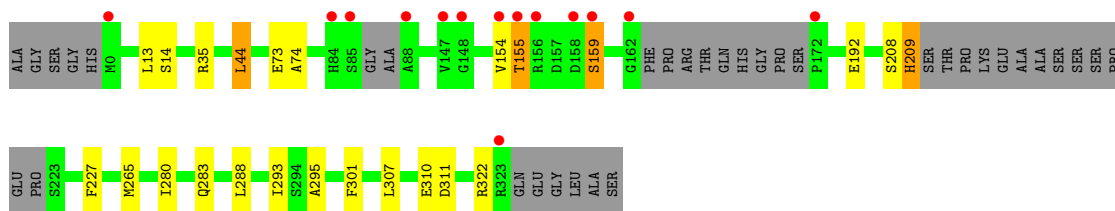


- Molecule 1: GATS-like protein 3



- Molecule 1: GATS-like protein 3





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	92.83Å 83.15Å 97.29Å 90.00° 116.39° 90.00°	Depositor
Resolution (Å)	33.03 – 2.05 33.03 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.4 (33.03-2.05) 94.3 (33.03-2.05)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.22 (at 2.05Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.185 , 0.211 0.188 , 0.214	Depositor DCC
R_{free} test set	4521 reflections (5.41%)	wwPDB-VP
Wilson B-factor (Å ²)	33.5	Xtrriage
Anisotropy	0.264	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 40.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.016 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	19732	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MLI

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.14	0/2447	0.36	0/3338
1	B	0.15	0/2399	0.36	1/3270 (0.0%)
1	C	0.16	0/2548	0.40	0/3476
1	D	0.15	0/2409	0.42	0/3280
All	All	0.15	0/9803	0.38	1/13364 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	153	PRO	N-CA-CB	6.77	110.44	103.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2383	2384	2380	5	0
1	B	2331	2335	2340	5	0
1	C	2479	2467	2465	8	0
1	D	2350	2345	2336	11	0
2	A	12	15	12	1	0
2	B	12	12	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	12	12	12	0	0
2	D	12	12	12	0	0
3	A	21	6	6	0	0
3	C	7	2	2	0	0
3	D	14	4	4	0	0
4	A	115	0	0	0	0
4	B	104	0	0	0	1
4	C	158	0	0	0	0
4	D	128	0	0	0	1
All	All	10138	9594	9581	29	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

The worst 5 of 29 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:227:PHE:CG	1:D:265:MET:HE1	2.41	0.56
1:A:187:LEU:HD13	1:A:254:LEU:HB3	1.92	0.51
1:C:280:ILE:O	1:C:283:GLN:HG2	2.13	0.48
1:D:155:THR:O	1:D:155:THR:HG23	2.13	0.48
1:B:13:LEU:N	1:B:13:LEU:HD23	2.29	0.47

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:600:HOH:O	4:D:623:HOH:O[1_655]	2.11	0.09

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	301/335 (90%)	294 (98%)	6 (2%)	1 (0%)	36	30
1	B	294/335 (88%)	286 (97%)	8 (3%)	0	100	100
1	C	315/335 (94%)	312 (99%)	3 (1%)	0	100	100
1	D	294/335 (88%)	283 (96%)	9 (3%)	2 (1%)	18	10
All	All	1204/1340 (90%)	1175 (98%)	26 (2%)	3 (0%)	36	37

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	155	THR
1	A	164	PRO
1	D	159	SER

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	264/289 (91%)	263 (100%)	1 (0%)	84	88
1	B	260/289 (90%)	255 (98%)	5 (2%)	50	50
1	C	277/289 (96%)	270 (98%)	7 (2%)	42	39
1	D	259/289 (90%)	252 (97%)	7 (3%)	39	36
All	All	1060/1156 (92%)	1040 (98%)	20 (2%)	50	50

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	35	ARG
1	D	192	GLU
1	D	322	ARG
1	D	209	HIS
1	C	13	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	302	ASN
1	D	108	HIS
1	D	135	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

10 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ARG	D	401	-	10,11,11	0.76	1 (10%)	9,13,13	0.97	1 (11%)
3	MLI	A	403	-	6,6,6	1.34	0	7,7,7	1.30	0
3	MLI	D	402	-	6,6,6	1.34	0	7,7,7	1.28	0
2	ARG	C	401	-	10,11,11	0.76	1 (10%)	9,13,13	0.95	1 (11%)
3	MLI	A	402	-	6,6,6	1.32	0	7,7,7	1.27	0
2	ARG	A	401	-	10,11,11	0.73	1 (10%)	9,13,13	0.90	1 (11%)
2	ARG	B	401	-	10,11,11	0.75	1 (10%)	9,13,13	0.95	1 (11%)
3	MLI	C	402	-	6,6,6	1.39	0	7,7,7	1.44	2 (28%)
3	MLI	D	403	-	6,6,6	1.39	0	7,7,7	1.48	2 (28%)
3	MLI	A	404	-	6,6,6	1.31	0	7,7,7	1.25	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ARG	D	401	-	-	1/11/11/11	-
3	MLI	A	403	-	-	4/4/4/4	-
3	MLI	D	402	-	-	2/4/4/4	-
2	ARG	C	401	-	-	1/11/11/11	-
3	MLI	A	402	-	-	0/4/4/4	-
2	ARG	A	401	-	-	0/11/11/11	-
2	ARG	B	401	-	-	1/11/11/11	-
3	MLI	C	402	-	-	0/4/4/4	-
3	MLI	D	403	-	-	0/4/4/4	-
3	MLI	A	404	-	-	2/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	401	ARG	OXT-C	-2.21	1.23	1.30
2	C	401	ARG	OXT-C	-2.20	1.23	1.30
2	A	401	ARG	OXT-C	-2.16	1.23	1.30
2	B	401	ARG	OXT-C	-2.14	1.23	1.30

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	ARG	OXT-C-O	-2.77	117.80	124.08
2	D	401	ARG	OXT-C-O	-2.70	117.95	124.08
2	C	401	ARG	OXT-C-O	-2.62	118.13	124.08
2	A	401	ARG	OXT-C-O	-2.57	118.26	124.08
3	D	403	MLI	O9-C3-C1	2.11	121.06	114.51

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	403	MLI	C2-C1-C3-O9
3	D	402	MLI	C3-C1-C2-O7
3	A	404	MLI	C3-C1-C2-O6
3	A	404	MLI	C3-C1-C2-O7
3	A	403	MLI	C2-C1-C3-O8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	ARG	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	305/335 (91%)	0.22	8 (2%) 57 59	26, 46, 78, 91	2 (0%)
1	B	296/335 (88%)	0.22	9 (3%) 52 53	21, 46, 81, 104	4 (1%)
1	C	318/335 (94%)	-0.04	16 (5%) 34 34	18, 36, 74, 105	3 (0%)
1	D	300/335 (89%)	0.14	14 (4%) 36 36	25, 42, 81, 97	2 (0%)
All	All	1219/1340 (90%)	0.13	47 (3%) 43 44	18, 43, 80, 105	11 (0%)

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	162	GLY	5.8
1	C	158	ASP	4.6
1	A	209	HIS	4.0
1	B	153	PRO	3.8
1	C	259	SER	3.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	MLI	D	402	7/7	0.76	0.14	47,54,58,62	0
3	MLI	D	403	7/7	0.79	0.12	53,57,73,73	0
3	MLI	A	402	7/7	0.81	0.13	60,64,79,84	0
3	MLI	A	404	7/7	0.81	0.12	40,56,66,67	0
3	MLI	A	403	7/7	0.85	0.11	56,68,78,80	0
3	MLI	C	402	7/7	0.90	0.12	41,46,65,68	0
2	ARG	B	401	12/12	0.94	0.07	28,34,41,41	0
2	ARG	A	401	12/12	0.96	0.06	27,32,36,38	0
2	ARG	D	401	12/12	0.96	0.06	27,32,39,39	0
2	ARG	C	401	12/12	0.97	0.05	18,25,30,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.